Spatio-Temporal Motion Estimation for Disease Discrimination in Cardiac Echo Videos

F Wang, T Syeda-Mahmood, D Beymer

IBM Almaden Research Center, San Jose, CA, USA

Abstract

In this paper we present a method of simultaneous registration of an entire sequence of frames of an echocardiographic sequence. In our approach, each echo frame is modeled using a probability density function, and registration problem between all pairs of echo frames is formulated as the problem of matching probability densities. An information-theoretic criterion called the Jensen-Renyi divergence is used to measure the distance between the probability density functions. The Renyi's Quadratic entropy results in a closed- form solution for the registration problem. Once the echo frames are registered, temporal trajectories of corresponding feature points in successive frames can be used to derive average velocity curves which have been shown to be useful for disease discrimination. To evaluate our technique for echo motion estimation for disease discrimination, we tested on a data set including cardiac echo from 21 patients of varying diseases. The data set includes a total of 72 complete cardiac cycles and contains 1612 frames. We compare our approach against two competing motion detection techniques, optical flow and Demons algorithm, on the same data set, and our motion detector performs best in terms of the separation between different diseases.

1. Introduction

Disease-specific understanding of echocardiographic sequences requires accurate characterization of spatio-temporal motion patterns. Similar motion patterns in the designated cardiac regions can signify similar disease for some patients. However, accurate characterization of these patterns requires registration of regions depicted in successive video frames. Previous methods of motion estimation have considered either successive pairs of echo frames or frames over a narrow time window. They exploit correspondence between points in successive frames using optical flow, or model motion estimation as the problem as a deformable surface fitting problem. To ensure consistency between temporal motion fields over longer range and better tracking, these methods often restrict the variations to lie along spline temporal models. However, due to the highly non-rigid nature of heart motion, these methods result frequently in incoherent motion estimates causing erroneous estimation of spatio-temporal motion patterns.

In this paper we present a method of simultaneous registration of an entire sequence of frames of an echocardiographic sequence. The approach uses probabilistic shape matching wherein each echo frame is modeled using a probability density function, and registration problem between all pairs of echo frames is formulated as the problem of matching probability densities. An new information-theoretic criterion called the Jensen-Renyi divergence is used to measure the distance between the probability density functions. The Renyi's Quadratic entropy results in a closed- form solution for the registration problem. Once the echo frames are registered, temporal trajectories of corresponding feature points in successive frames can be used to derive average velocity curves which have been shown to be useful for disease discrimination. To evaluate our technique for echo motion estimation for disease discrimination, we tested on a data set including cardiac echo from 21 patients of varying diseases. The data set includes a total of 72 complete cardiac cycles and contains 1612 frames. We compare our approach against two competing motion detection techniques, optical flow and Demons algorithm, on the same data set, and our motion detector performs best in terms of the separation between different diseases.

The estimation of cardiac motion constitutes an important aid for the quantification of the elasticity and contractility of the myocardium. Motion estimation techniques that are based on a closed form solution to the registration problem not only give more accurate registration but also lead to the extraction of useful features for disease discrimination and decision support.

The rest of the paper describes our approach in detail. In the rest of this section, we will review related work on motion estimation for automatic cardiac disease discrimination of cardiac echo videos. In Section 2, we describe our overall approach to disease recognition from spatio-temporal models. Finally, in Section 3, we present results of disease discrimination of hypokinesia patients over nor-

mal patients.

1.1. Related work

The estimation of cardiac motion and deformation from cardiac imaging has important clinical implications for assessment of viability in the heart wall. Different approaches have been proposed for motion recovery from 2D echocardiographic video sequences. The most popular approach is the optical flow based methods[1,2], however the estimates are often inconsistent with the actual observed motion in the echo videos for cardiac regions due to the low quality of the echo videos and non-smooth heart motion. Another approach is myocardial region segmentation using deformable models [3], and the motion is then recovered by aligning the segmented shapes. The dependency on obtaining an accurate segmentation, however, remains a significant issue, as there still are no fully automated robust and efficient. In Papademetris et al. [4], an Bayesian approach, combined with the biomechanical model were used the recover left ventricular deformation. It has the advantage of accounting for the fiber directions in the left ventricle, however this approach also depends on a good segmentation of the heart. Finally, B-Spline based motion estimation techniques [5] often restrict the variations to lie along spline temporal models, which could leads to inaccurate motion estimations when the movement regions are far away from the spline control points.

While there is considerable work in cardiac echo region motion estimation, not much work exists for automatic disease discrimination. Recently, there have been efforts to discriminate diseases by analyzing spatio-temporal properties of heart regions. An approach for validating disease diagnosis through video similarity was reported in [6], which used features from the entire heart region, restricting its use in characterizing region-specific diseases. Later work combined motion estimation using the Demon's algorithm with graph-based region segmentation approach [7] to improve disease discrimination. Even so, the inaccuracies of region segmentation using a graph-theoretic approach and rough motion estimation using average velocities often lead to inaccuracies in disease discrimination.

2. Methods

We now present the details of our proposed spatio-temporal motion estimation for disease discrimination algorithm. Our approach is based on a key observation that similar motion patterns in the designated cardiac regions can signify similar disease for some patients. Accurate characterization of these patterns requires registration of regions depicted in successive video frames. Therefore a consistent and accurate motion estimation is required. The overall approach is as follows. For each disease and each

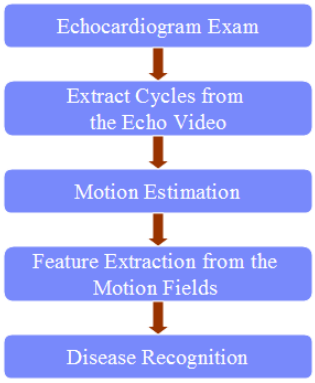

Figure 1. Illustration of the processing steps of our ECG and heart sound based disease similarity search algorithm.

of its diagnostic viewpoints, we first extract cycles from the echo video with the help of the ecg embedded in the video; We then use the Jensen-Renyi divergence to achieve the alignment of all frames within a heart cycle; Finally motion features such as "average velocity curve" are used to discriminate different diseases. Each step of our algorithm is illustrated in Fig. 1. We begin with the motion estimation module, which align all the echo frames within a heart cycle.

2.1. Motion estimation

In this section, we formulate the groupwise motion estimation using JR divergence. Let $I_1(x), I_2(x), \ldots, I_n(x)$ be the successive frames that we want to estimate the motion from, and they are within one cycle of the heart beat. We represented these intensity images as probability density functions. We then pose the simultaneous alignment of multiple frames as the problem of determining the transformation parameters μ^i such that the Jensen-Renyi divergence between the probability density functions of the echocardiographic images is minimized.

The Jensen-Renyi divergence [8] between probability density functions is defined as:

$$JR_\beta(P_1, P_2, ..., P_N) = H_\alpha(\sum \beta_i P_i) - \sum \beta_i H_\alpha(P_i) \quad (1)$$

where $H_\alpha(X)$ is Renyi entropy of order α ($\alpha \geq 0$), which is defined on a random variable X and is given by

$$H_\alpha(X) = \frac{1}{1-\alpha} \log(\sum_{i=1}^{n} p_i^\alpha) \quad (2)$$

where P_i are the probabilities of echocardiographic frames $I_1(f_1(x;\mu^1), I_2(f_2(x;\mu^2), ..., I_n(f_n(x;\mu^n)$ within a single heart cycle, and the transformation $f_i(x;\mu^i)$ describes the deformation from the I_i to the reference frame, which is chosen to be the frame corresponding to the ECG R peak

(beginning of the systole). We use the Parzen window technique to estimate the probability density function of an image, if the kernel of the Parzen window is chosen to be a Gaussian, the P_i can be viewed as a Gaussian Mixture model.

The groupwise registration problem can then be formulated as the problem of minimizing

$$arg\min_{\mu^i} JR_\beta(P_1,P_2,...,P_N) + \lambda \sum_{i=1}^{N} ||Lf_i||^2 \qquad (3)$$

where Lf^i is a regularization term to control the nature of deformation. Note in Eqn. (2) above, the Renyi entropy is a generalization of Shannon entropy, since $\lim_{\alpha->1} H_\alpha = H_S$. Thus Jensen-Shannon divergence [9] is a special case of Jensen-Renyi divergence when $\alpha \to 1$. When $\alpha = 2$, $H_2 = -\log(\sum_{i=1}^{n} p_i^2)$ is called quadratic entropy. We choose α to be 2 in our implementation of JR divergence to estimate the motion, since it resulted in a closed-form expression of the JR divergence if the probability density function are represented as Gaussian Mixtures. We also derive the analytic gradient of this match measure in order to achieve efficient and accurate non-rigid registration. The details of the derivation is quite involved, and will be omitted here because of the space limit. The Jensen-Renyi measure is then minimized over a class of smooth non-rigid transformations, and the unknown displacement field are updated iteratively until converges. Our motion estimation is quite distinct from those existing in literature because we are using a new information-theoretic criterion, and it is achieving the non-rigid registration simultaneously instead done frame by frame.

2.2. Disease recognition using motion features

To capture the overall motion of the heart region better, we use the average velocity curve described in [6]. The average velocity curve preserves a common sense of perceived motion per direction and is obtained by averaging the speed and direction of the velocity vectors at each pixel in the region within each frame.

If we denote the velocity vectors per pixel (i, j) within the region of interest in frame k as (u_{ij}, v_{ij}), the average velocity vector in frame k is given by $(\delta_{avg}, \theta_{avg})$, where

$$\begin{aligned} \delta_{avg} &= \frac{\sum_i \sum_j \sqrt{(u_{ij}^2 + v_{ij}^2)}}{N} \\ \theta_{avg} &= \frac{\sum_i \sum_j tan^{-1} \frac{u_{ij}}{v_{ij}}}{N} \end{aligned} \qquad (4)$$

where N is the number of image pixels in the region. The average velocity curve is then given by $(C(t) = (x(t), y(t))$ where

$$\begin{aligned} x(t+1) &= x(t) + \delta_{avg} cos\theta_{avg} \\ y(t+1) &= y(t) + \delta_{avg} sin\theta_{avg} \end{aligned} \qquad (5)$$

By taking the projection of the average velocity curve along x, y, and t, we can obtain three additional region features. In particular, the projection onto x, y gives the total extent of planar motion of the region and is a good indication of the mechanical performance of the corresponding anatomical region. We also measure the area inside the region surrounded by the projection of the average velocity curve, which has been shown to be a good feature to discriminate between normal and abnormal echo videos [7].

To better capture the variance of the estimated motion field, one other motion feature that we use is the histogram of the motion field, whereas we calculate the magnitude and orientation histogram respectively with 18 bins for each histogram. Therefore each bin covers 20 degrees for the orientation histogram.

The final feature set will include the average velocity curve, area enclosed by the projection of the AVC, and motion histograms. These features will be then fed to a machine learning technique, where we chose to use the standard support vector machine (SVM) method for its proven ability to classify high dimensional nonlinear data. A subsets of the patient data sets with known disease labels are as training data, and the accuracy of our algorithm will be tested on the rest of the data sets.

3. Results

We now present experimental results on a 2 class problem: hypokinesia ($\mathbf{C}_0$) versus normal patients ($\mathbf{C}_1$). Hypokinesia is a cardiac disease where the heart suffers from reduced motion, so we expect the motion models to be quite different between the two classes. Fig. 2 shows the mean motion of normal patients versus Hypokinesia patients. It can be seen from this figure that the motion of normal patients is much greater than the hypokinetic patients. Thus, we expect the motion features to differentiate hypokinesia from normal patients.

Our data set came from hospitals in India with cardiologists recording a complete echo exam in continuous video. The diagnostic viewpoint we use for hypokinesia is Apical 4 Chamber (A4C). From these videos of the complete workflow, we manually extracted individual A4C cardiac cycles from both normal and hypokinesia patients. The echo "workflow" video was captured at 320x240 and 25 Hz, and a ECG trace waveform at the bottom allowed us to synchronize extracted cycles with the ECG R peak. Each video sequence was about 5 minutes long per patient and depicted several heart cycles. Our current collection has over 200 echo videos or 200x5x30x60 frames. The physicians in our project assisted us in their interpretation, so

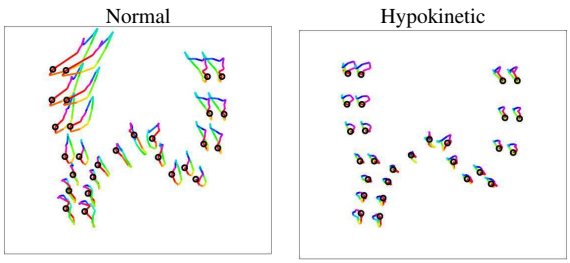

Figure 2. Motion pattern for normal patients and hypokinesia patients. (Best viewed in color.)

Recog. Rate	All patients (N = 21)	Hypokinesia $\mathbf{C}_0$ (N = 16)	Normal $\mathbf{C}_1$ (N = 5)
JR Div.	76.19%	75.0%	80.0%
Demons	61.90%	62.5%	60.0%
Optical Flow	57.14%	56.3%	60.0%

Table 1. Comparison of the disease recognition rate for different motion estimation methods.

that the ground truth labels for the cardiac cycles could be obtained. Due to the effort in manual labeling for model training, we report our results on a data set with 16 hypokinesia patients and 5 normal patients.

The testing results for our disease recognition algorithm are reported in Table 1 together with the results generated using Demons algorithm and standard optical flow approach. During testing, because of our data set size, we employed a leave-one-out methodology: when testing a sequence from patient X, the entire data set minus X is used for training. From Table 1, it is evident that our method achieve a better recognition rate than the other two rival techniques. We believe that increasing the number of patients will lead to better recognition rates for all the techniques.

4. Discussion and conclusions

In this paper we have demonstrated that estimates of cardiac deformation can be obtained from ultrasound images using an information theoretic approach. Our method is able to simultaneous register an entire sequence of frames of an echocardiographic sequence. Once the echo frames are registered, motion features such as temporal trajectories of corresponding feature points in successive frames can be used to derive average velocity curves which have been shown to be useful for disease discrimination. Our motion estimation technique performs best in terms of the separation between different diseases when compared against two competing techniques. Future work will involve more extensive experiments on more data sets from a large number of disease classes.

References

[1] Behbahani S, Magholi K. Evaluation optical-flow based methods for estimation of wallmotions. In Second International Multi-Symposiums on Computer and Computational Sciences. IEEE Computer Society, 2007; 164–169.

[2] Macan T, Loncaric S. 3-d cardiac motion estimation by point-constrained optical flow algorithm, 2001.

[3] Jacob G, Noble JA, Behrenbruch CP, Kelion AD, Banning AP. A shape-space based approach to tracking myocardial borders and quantifying regional left ventricular function applied in echocardiography. IEEE Trans Med Imaging 2002; 21(3):226–238.

[4] Papademetris X, Sinusas AJ, Dione DP, Duncan JS. Estimation of 3D left ventricular deformation from echocardiography. Medical Image Analysis 2001;17–28.

[5] Ledesma-Carbayo M, Kybic J, Desco M, Santos A, Suhling M, Hunziker P, Unser M. Spatio-Temporal Nonrigid Registration for Ultrasound Cardiac Motion Estimation. IEEE Transactions on Medical Imaging 2005;24(9):1113–1126.

[6] Syeda-Mahmood TF, Yang J. Characterizing normal and abnormal cardiac echo motion patterns. In Computers in Cardiology, volume 33. 2006; 725–728.

[7] Syeda-Mahmood T, Wang F, Beymer D, London M, Reddy R. Characterizing spatio-temporal patterns for disease discrimination in cardiac echo videos. In MICCAI. 2007; 261–269.

[8] Hamza AB, Krim H. Jensen-renyi divergence measure: theoretical and computational perspectives. IEEE International Symposium on Information Theory 2003;257–257.

[9] Lin J. Divergence measures based on the Shannon entropy. IEEE Trans Infor Theory 1991;37:145–151.

Address for correspondence:

Fei Wang
650 Harry Road, San Jose, CA 95120
wangfe@us.ibm.com